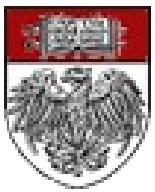
Hidden Valley Higgs Search
Update

Higgs Group

Shawn Kwang
Mel Shochet

University of Chicago

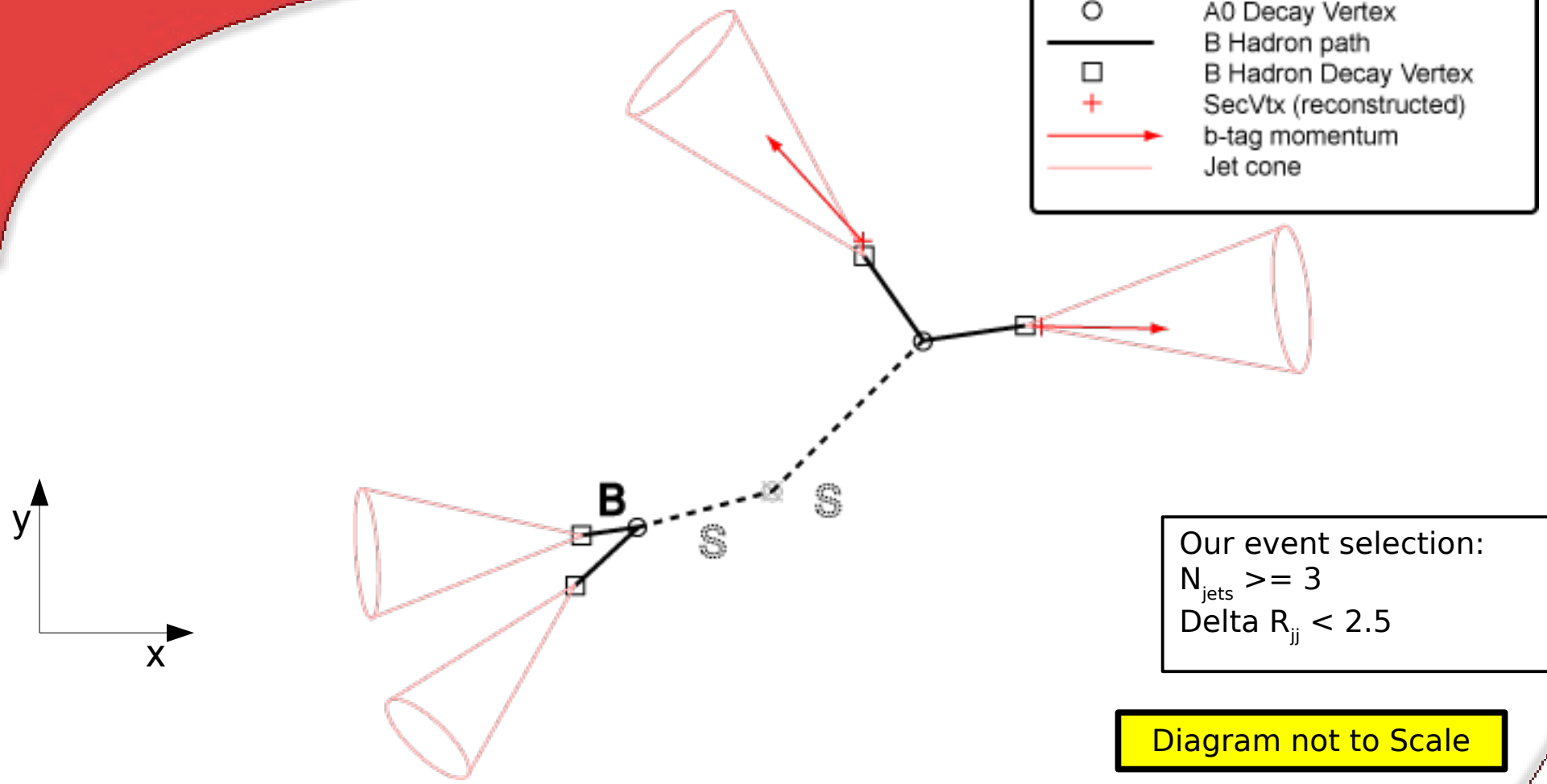


Introduction

- I described the algorithm to generate the background estimate (from data) in my previous presentation: 2010-04-23 Higgs Group Meeting.
- Lots of questions were asked, and in this update I present answers.

- Review of the analysis
  - Signal we are searching for
  - Event Selection
  - Algorithm
  - Discriminant Variables
- Questions and Answers
- Changes to the algorithm resulting from feedback

Model Diagrams



Our event selection:
N_{jets} >= 3
Delta R_{jj} < 2.5

Diagram not to Scale

Here the Higgs decays at the primary vertex (the X). S represents the heavy pseudoscalar with a long lifetime, which decays into bbbar pairs.

The pink cones represent the hadronization of the B hadrons into jets.

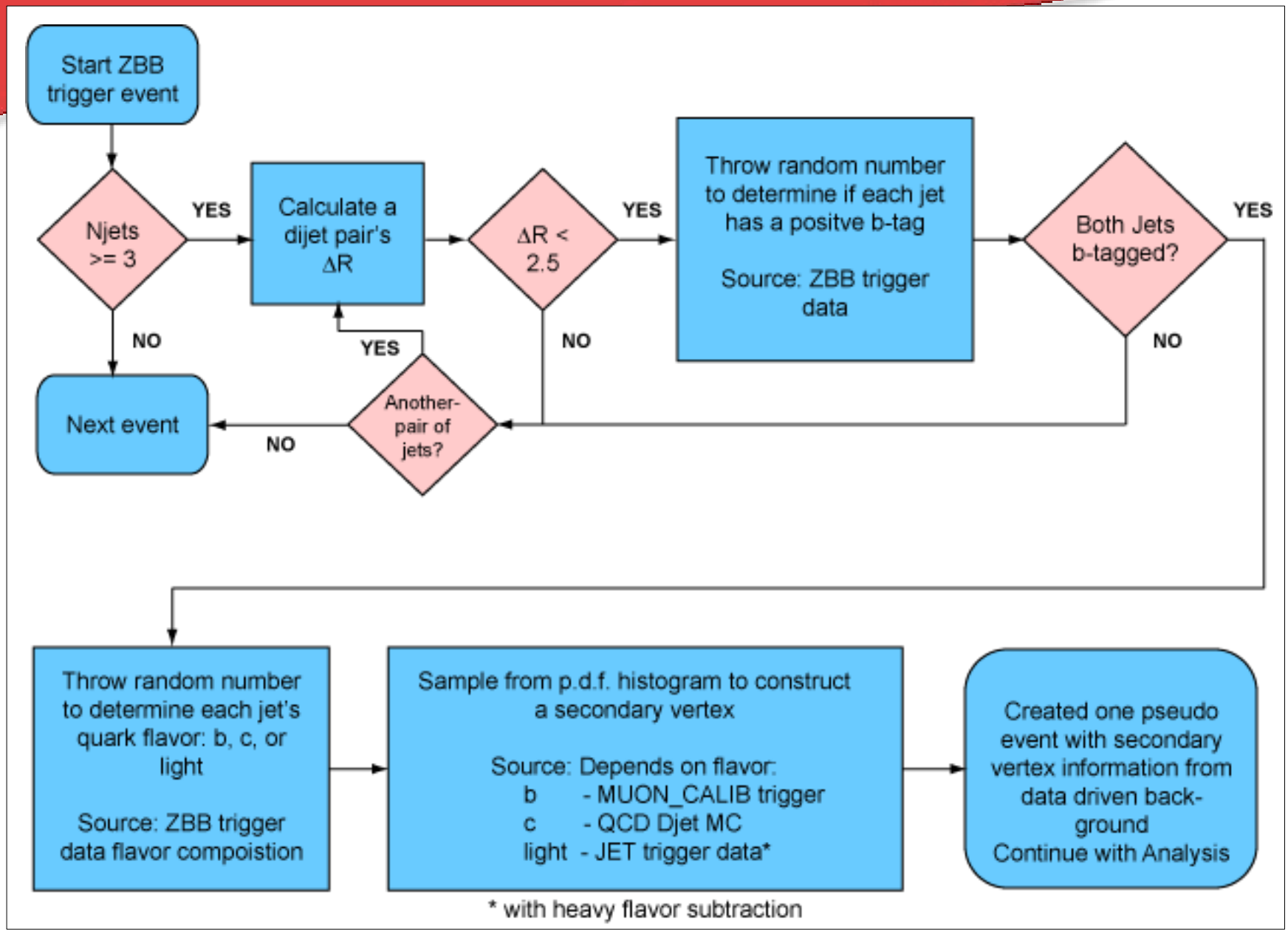
The red represents reconstructed secondary verticies and their corresponding momenta.

Black is the “truth” information.

Signal Event Selection

- We will be looking for events with central b-tagged jets, with a relatively low E_T requirement. "Signal Region"
  - All jets are required to have:
    - E_T > 20 GeV, corrected at Level-5
    - $|\eta|<1.0$
  - Jet multiplicity: $N_{jet} \geq 3$
  - For the dijet system, require that it be in a region that would be populated by signal.
    - $\Delta R < 2.5$
- A "Control Region" is defined which contains events orthogonal to the Signal Region,
  - Two tight central jets ($N_{jet} = 2$)
  - A third jet with Level 5 corrected $E_T < 15$ GeV.

Flowchart



This flowchart shows the background estimate / pseudo event generating algorithm.
It is here for reference.

Diagrams ψ

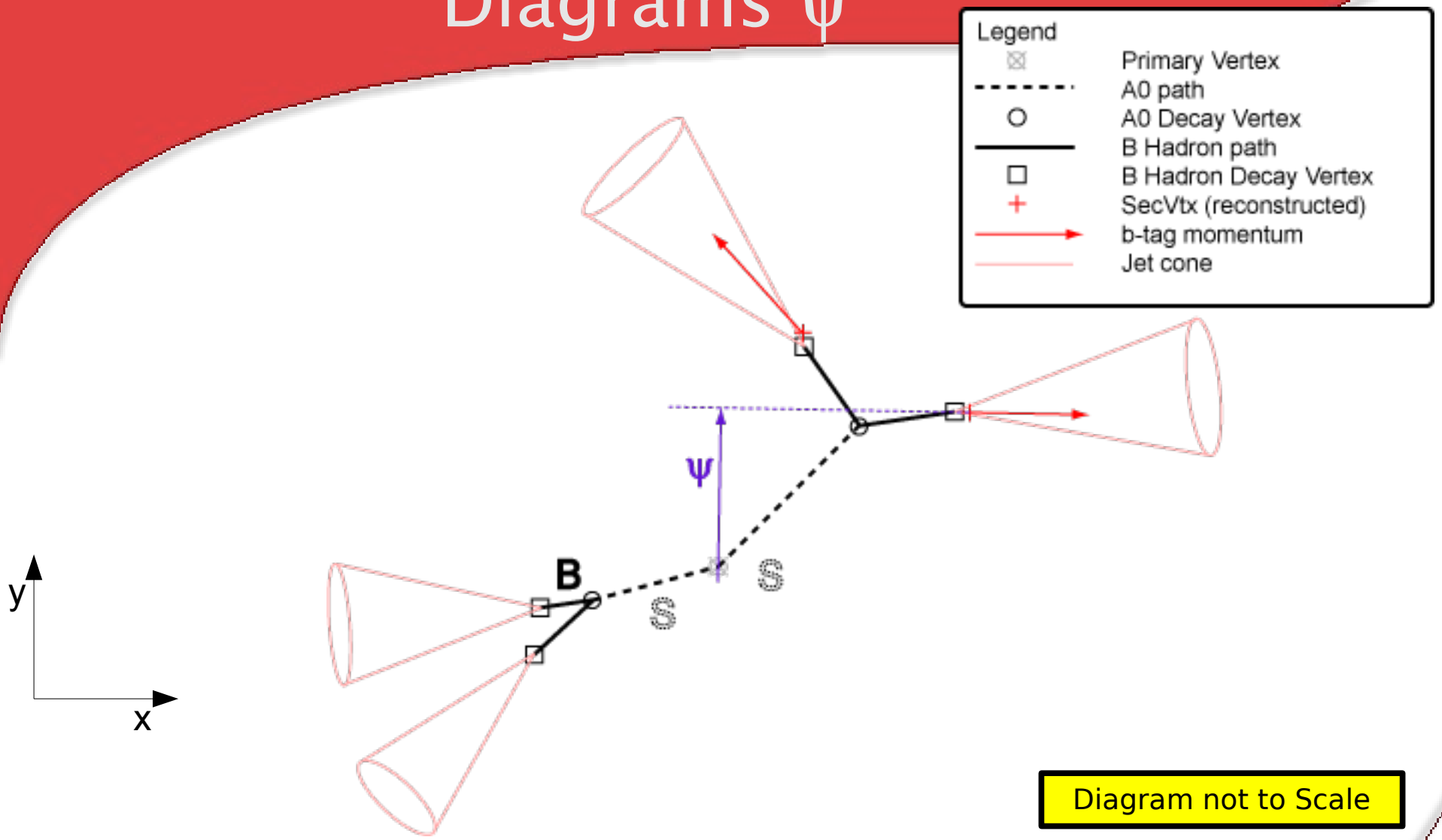


Diagram not to Scale

ψ is the impact parameter of a jet with a secondary vertex.

This is in two-dimensional space.

Diagrams ζ

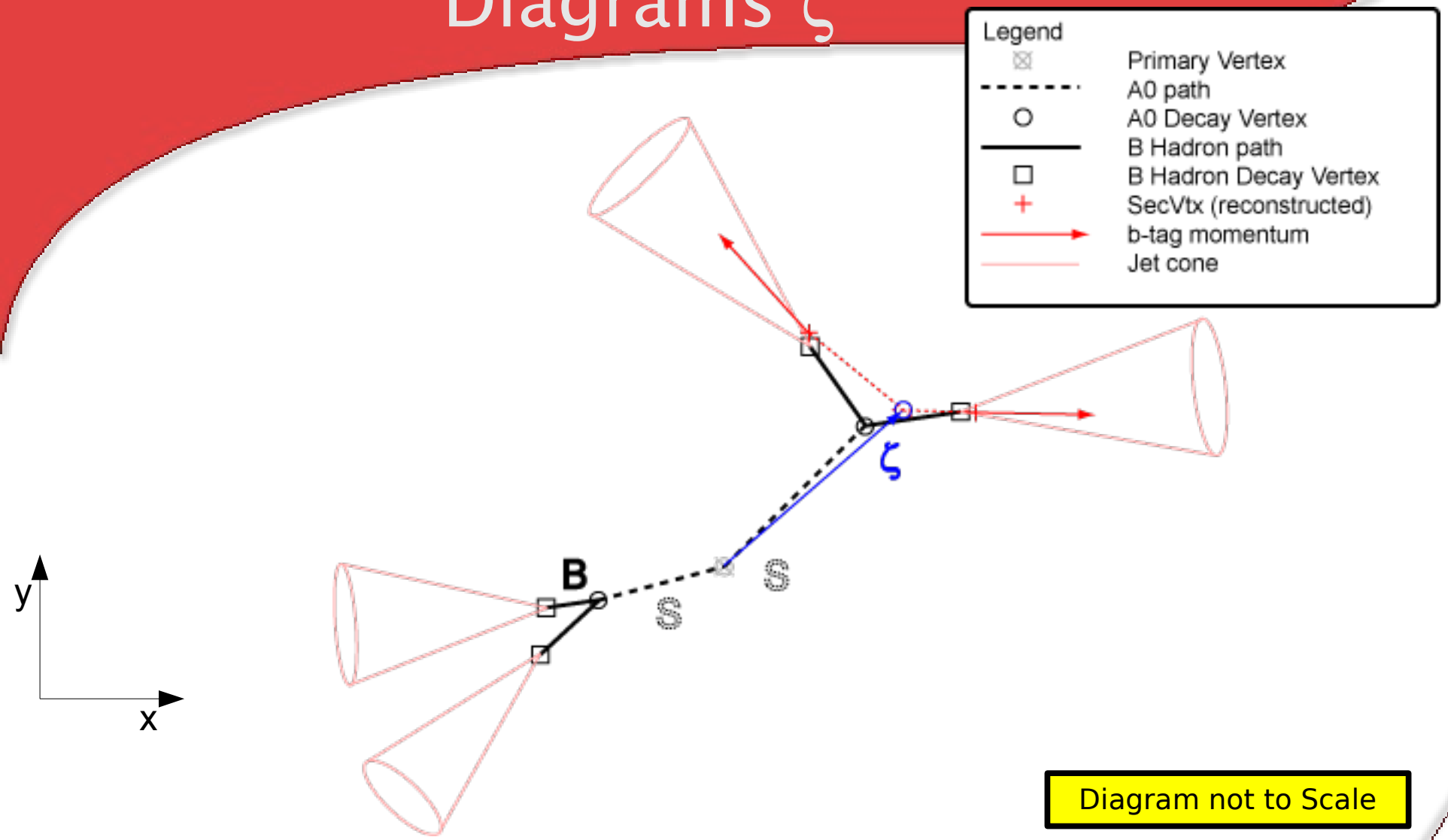


Diagram not to Scale

ζ is the reconstructed decay distance of the heavy pseudoscalar S (A_0). It requires two tagged jets.

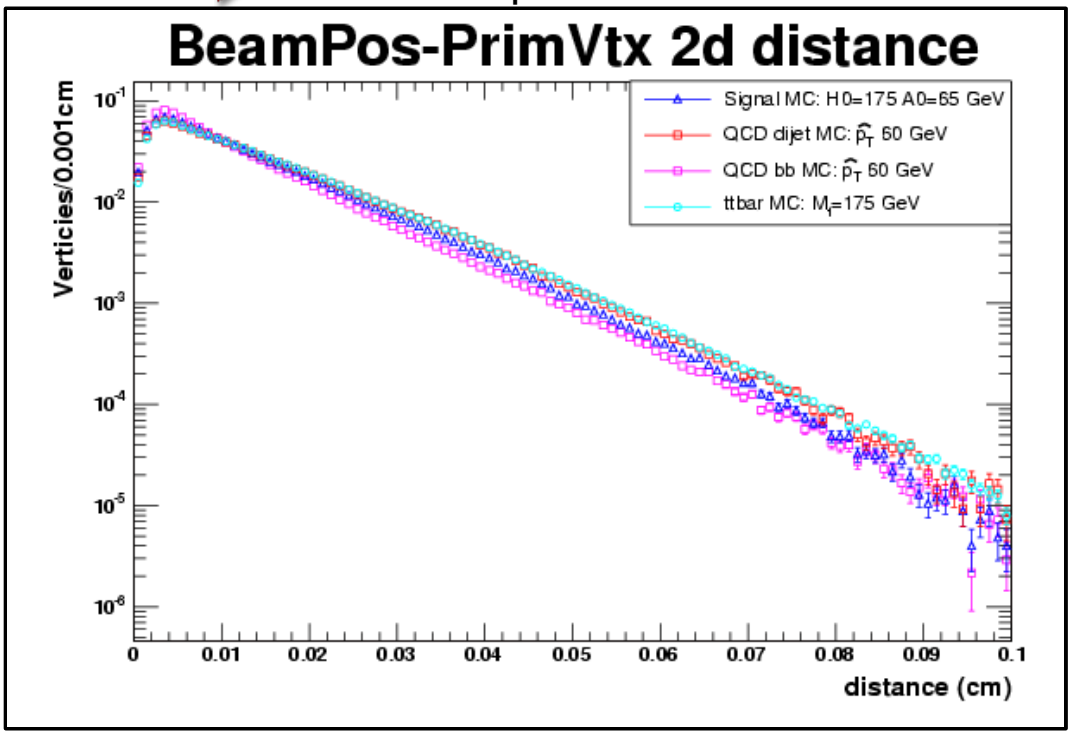
This is in two-dimensional space.

Questions

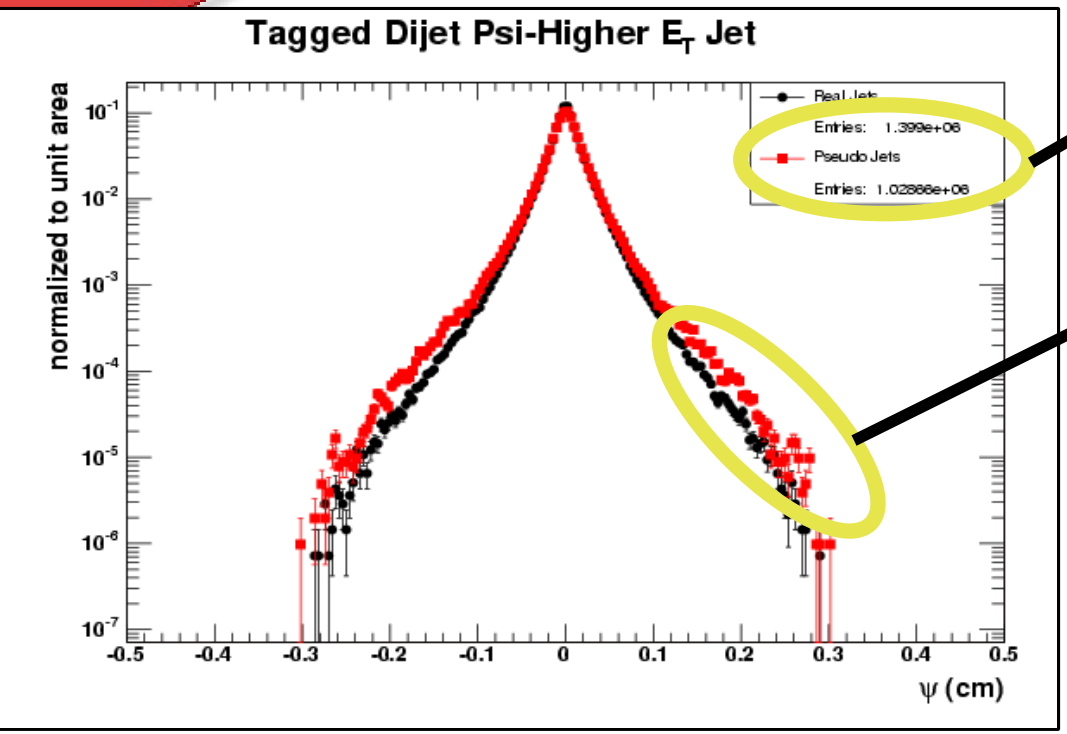
- Does this signal topology bias the Primary Vertex?
- What is the origin of the rate and shape differences in the Control region, seen in variable ψ.
- Dijet correlations: Whether or not the correlations between two jets affect the analysis.
  - “Does is make sense to assume that the kinematics of the two jets are independent when calculating Zeta?”
  - “Does it make sense to assume the flavor of the two jets in event are not correlated ?”

Primary Vertex

- "Check in signal MC how often you get the primary vertex correct."
- The signal we are searching for has the topology where two long lived particles decay away from the primary vertex. Does the primary vertex reconstruction accurately find the real primary or is there a bias?
- I plot the distance from the beam position to the primary vertex in 2D space.
  - Signal MC, QCD bb MC, QCD dijet MC, and top MC
  - No significant difference is seen between these MC samples
- The beam position is obtained from the database in these MC samples.



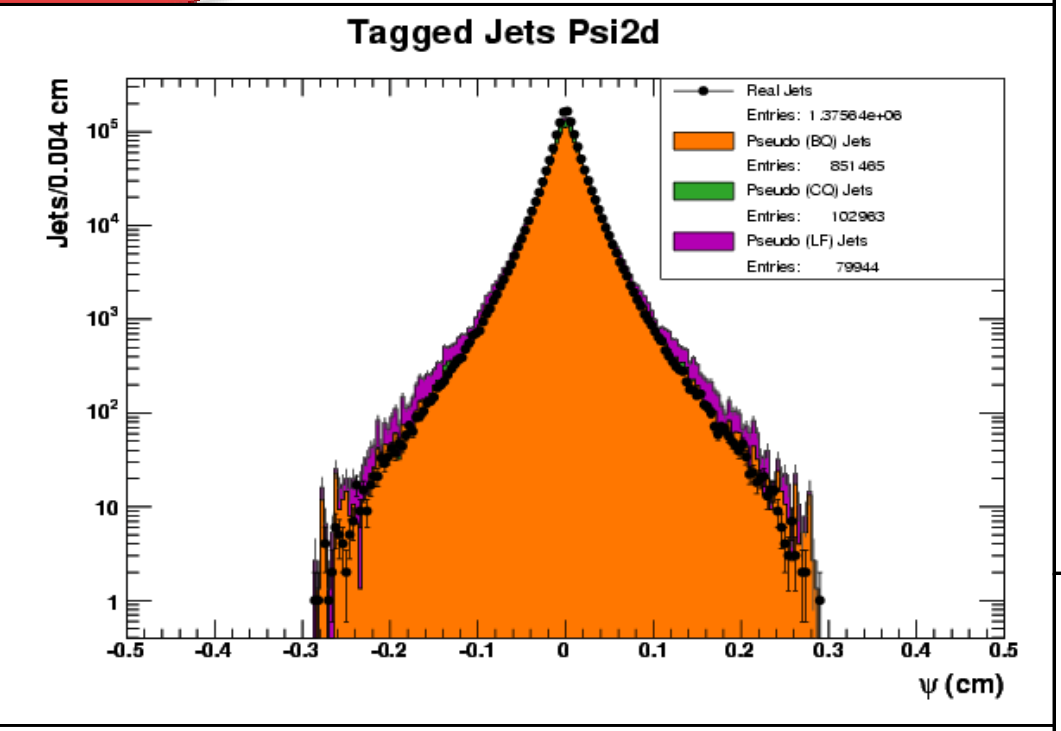
Questions regarding ψ

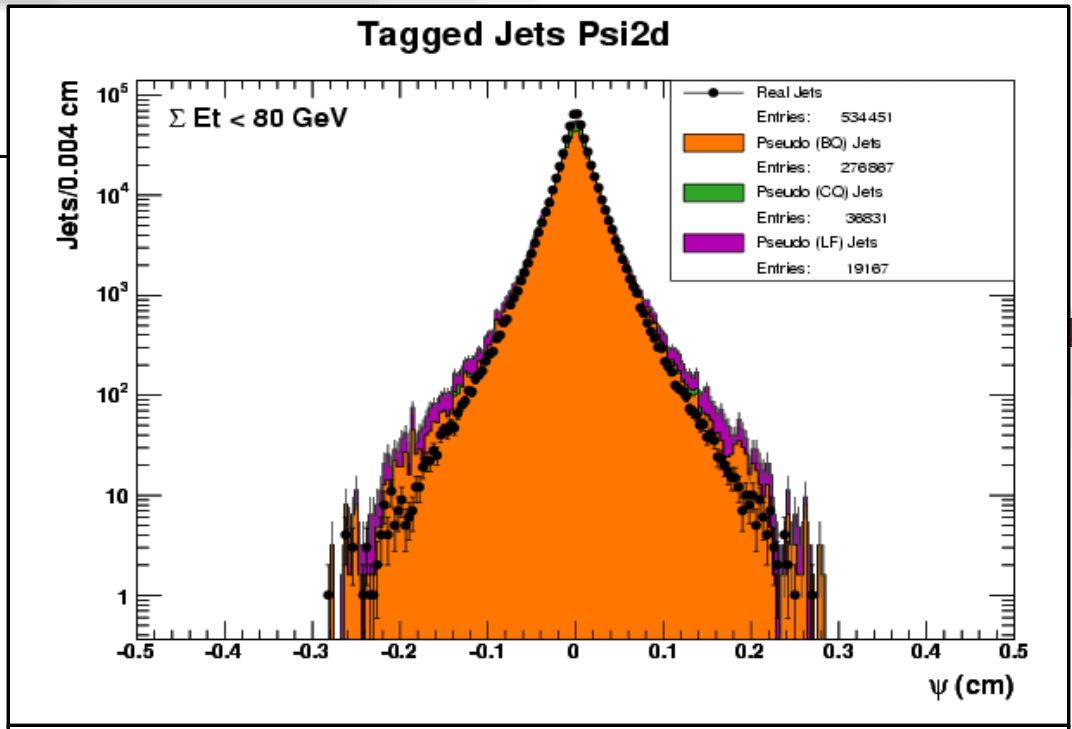


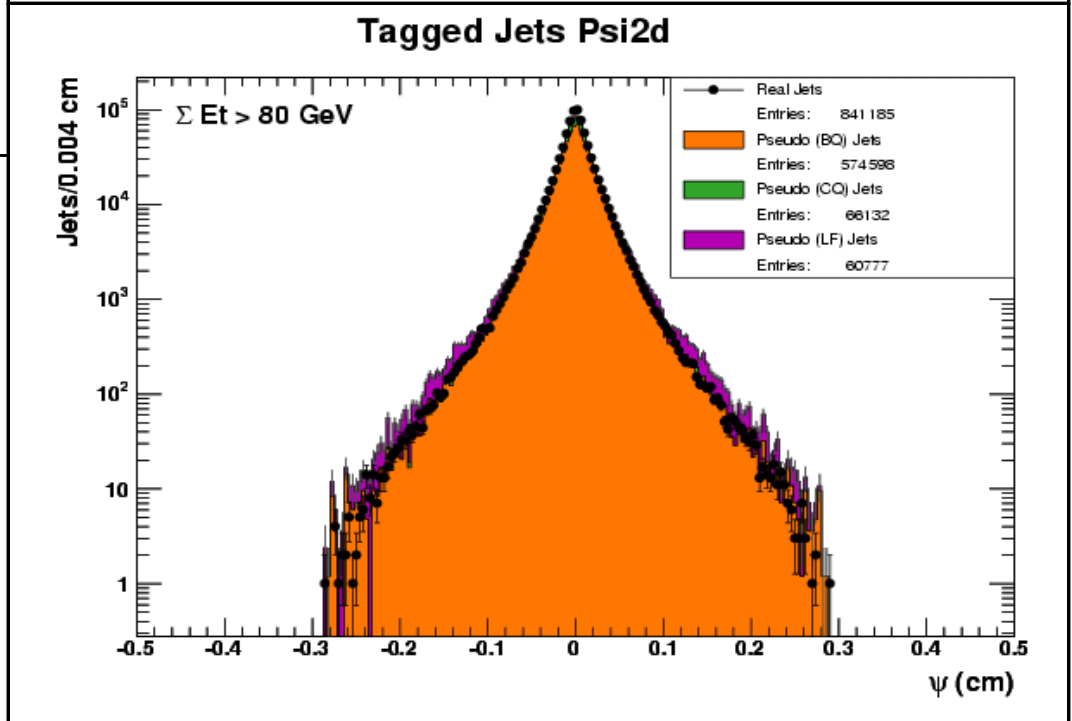
- Number of Entries are different
  - Real: 1.399M
  - Pseudo: 1.026M
- Shapes are different when normalized

- What is plotted are all events from all bins for a particular max d_0 cut ($|d_0| < 0.30$ cm).
- The plot on the resulted in many questions, specifically with the differences in the shapes of the real and pseudo data, as well as the rates.

ψ Histograms







- What is plotted are all events from all bins for a particular max d_0 cut ($|d_0| < 0.30$ cm).
- The pseudo data has been split into different quark flavors and stacked.
- On the right are two distributions for different event ΣE_T.

ψ Shape differences

- In addition, I looked at ψ in other variables, such as eta, and found that there is no difference.
- In the end this shape difference may have to be taken as a systematic.
- I remind you that if we are overestimating the background rate along the tails, then our background will be too large, which is preferable to the opposite.

Rate differences

- We see a difference in the rates between real and pseudo jets.
  - In both the 2 jet control and 3 jet signal regions.

Et Bin	SVT Bin	Ctrl real jets	Ctrl pseudo jet	% difference
20-30	SVTTRK 0	206479	145028	-42.37
20-30	SVTTRK 1	545330	383515	-42.19
20-30	SVTTRK ≥ 2	481443	381440	-26.22
30-70	SVTTRK 0	201844	157128	-28.46
30-70	SVTTRK 1	588806	430436	-36.79
30-70	SVTTRK ≥ 2	663644	518471	-28.00
70-110	SVTTRK 0	8880	8434	-5.29
70-110	**SVTTRK 1**	**21820**	**18404**	**-18.56**
70-110	SVTTRK ≥ 2	25764	20361	-26.54
110-200	SVTTRK 0	1312	1157	-13.40
110-200	SVTTRK 1	2868	2175	-31.86
110-200	SVTTRK ≥ 2	3082	2195	-40.41
	Total	2751272	2068744	-32.99

Et Bin	SVT Bin	Signal real jets	Signal pseudo jet	% difference
20-30	SVTTRK 0	28743	32937	12.73
20-30	SVTTRK 1	42087	45759	8.02
20-30	SVTTRK ≥ 2	41729	44391	6.00
30-70	SVTTRK 0	25263	27478	8.06
30-70	SVTTRK 1	42031	41434	-1.44
30-70	SVTTRK ≥ 2	54888	53442	-2.71
70-110	SVTTRK 0	2041	2612	21.86
70-110	**SVTTRK 1**	**3377**	**3629**	**6.94**
70-110	SVTTRK ≥ 2	4536	4276	-6.08
110-200	SVTTRK 0	391	607	35.58
110-200	SVTTRK 1	620	778	20.31
110-200	SVTTRK ≥ 2	720	803	10.34
	Total	246426	258146	4.54

B−tag Probability

Et Bin	SVT Bin	Probability any jet is tagged(%)	Control-probability 2nd jet is tagged given the first is tagged(%)
20-30	SVTTRK 0	4.24	5.69
20-30	SVTTRK 1	22.95	33.64
20-30	SVTTRK ≥ 2	64.99	76.16
30-70	SVTTRK 0	5.48	6.91
30-70	SVTTRK 1	20.71	31.21
30-70	SVTTRK ≥ 2	59.04	70.46
70-110	SVTTRK 0	6.06	6.75
70-110	**SVTTRK 1**	**13.6**	**18.99**
70-110	SVTTRK ≥ 2	34.27	45.33
110-200	SVTTRK 0	6.41	6.76
110-200	SVTTRK 1	11.25	14.42
110-200	SVTTRK ≥ 2	22.99	30.54

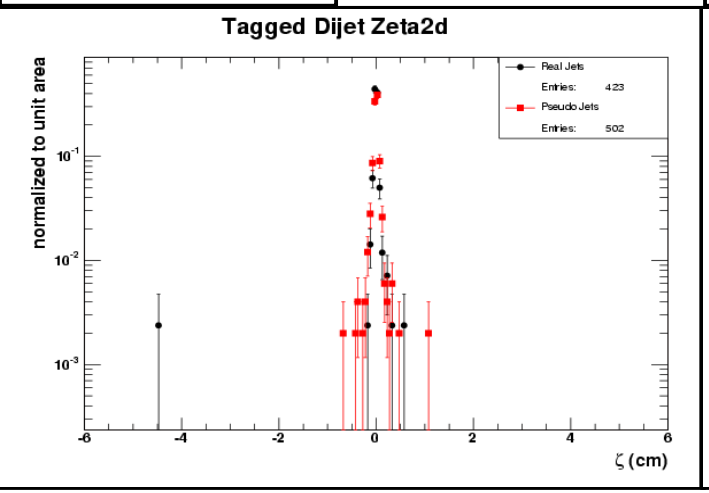
- The first column is that probability that any jet has a b-tag
- The second column is conditional, the first jet (of the dijet pair) must already have a b-tag, then the probability of the second jet having a b-tag is calculated.
- There is a correlations between b-tags in the dijet pair.
- This will have to be taken into account in the new pseudo event generation algorithm.

Kinematic Correlations of Dijets

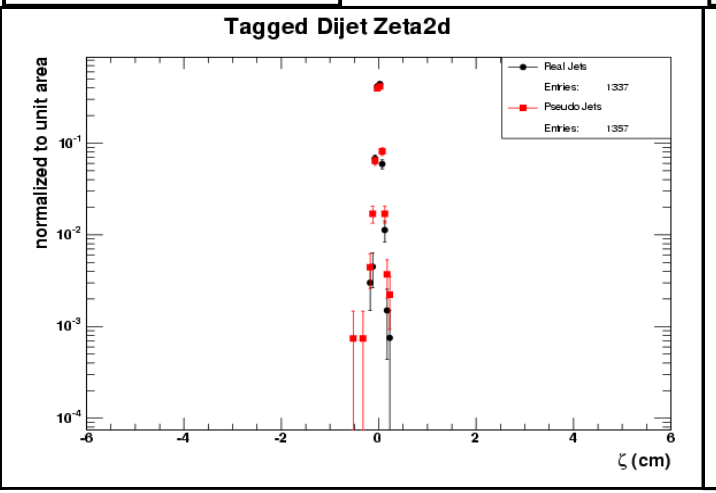
- “Does is make sense to assume that the kinematics of the two jets are independent when calculating Zeta?”
- What we are doing is selecting ZBB trigger events and using the kinematic information of that event as is.
- Then we select where the b-tag/secondary vertex exists in space, in a similar way nature does it.
  - Nature should decide the decay vertex independently of each other.
  - Thus the position of one secondary vertex should not depend on another.
  - As a result the ζ variable should not be dependent on the kinematics of the two jets.
- On the next slide I show ζ for different ΔR bins.

ζ Histograms

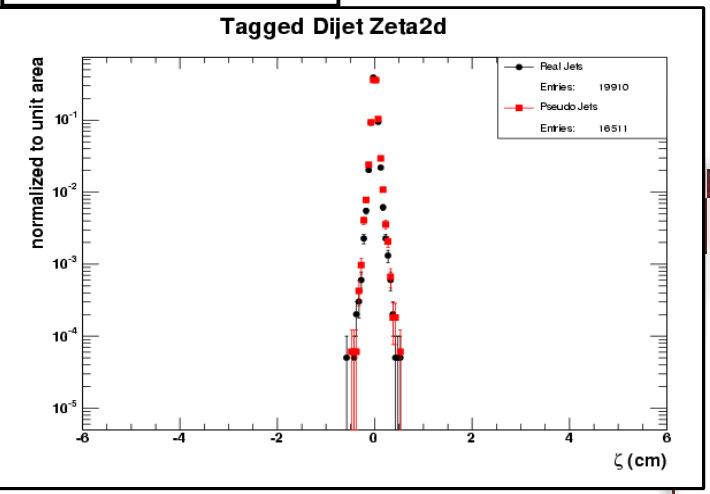
$1.0 \leq \Delta R < 1.5$

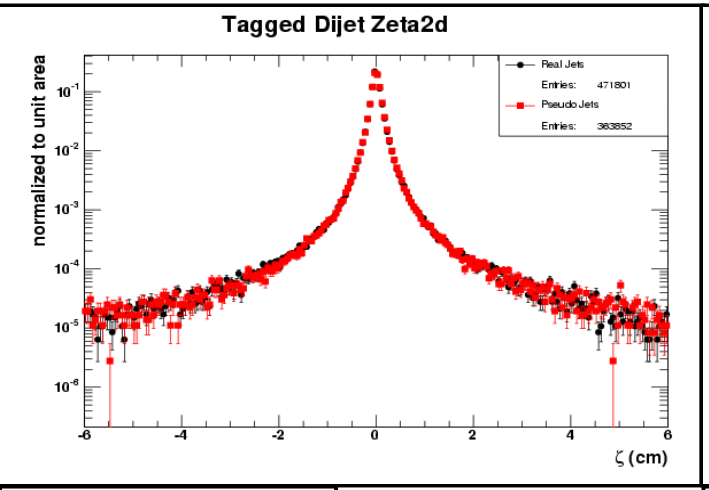


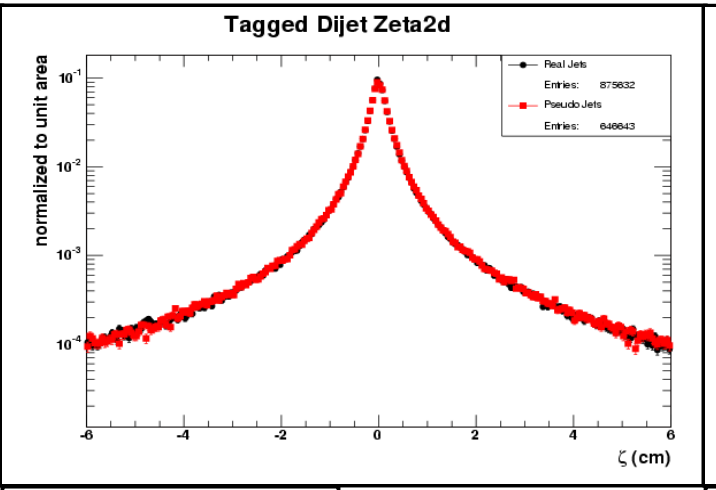
$1.5 \leq \Delta R < 2.0$

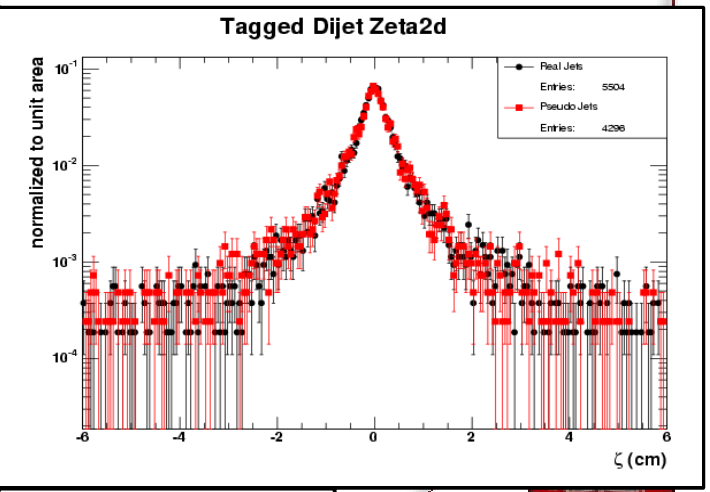


$2.0 \leq \Delta R < 2.5$









$2.5 \leq \Delta R < 3.0$

$3.0 \leq \Delta R < 3.5$

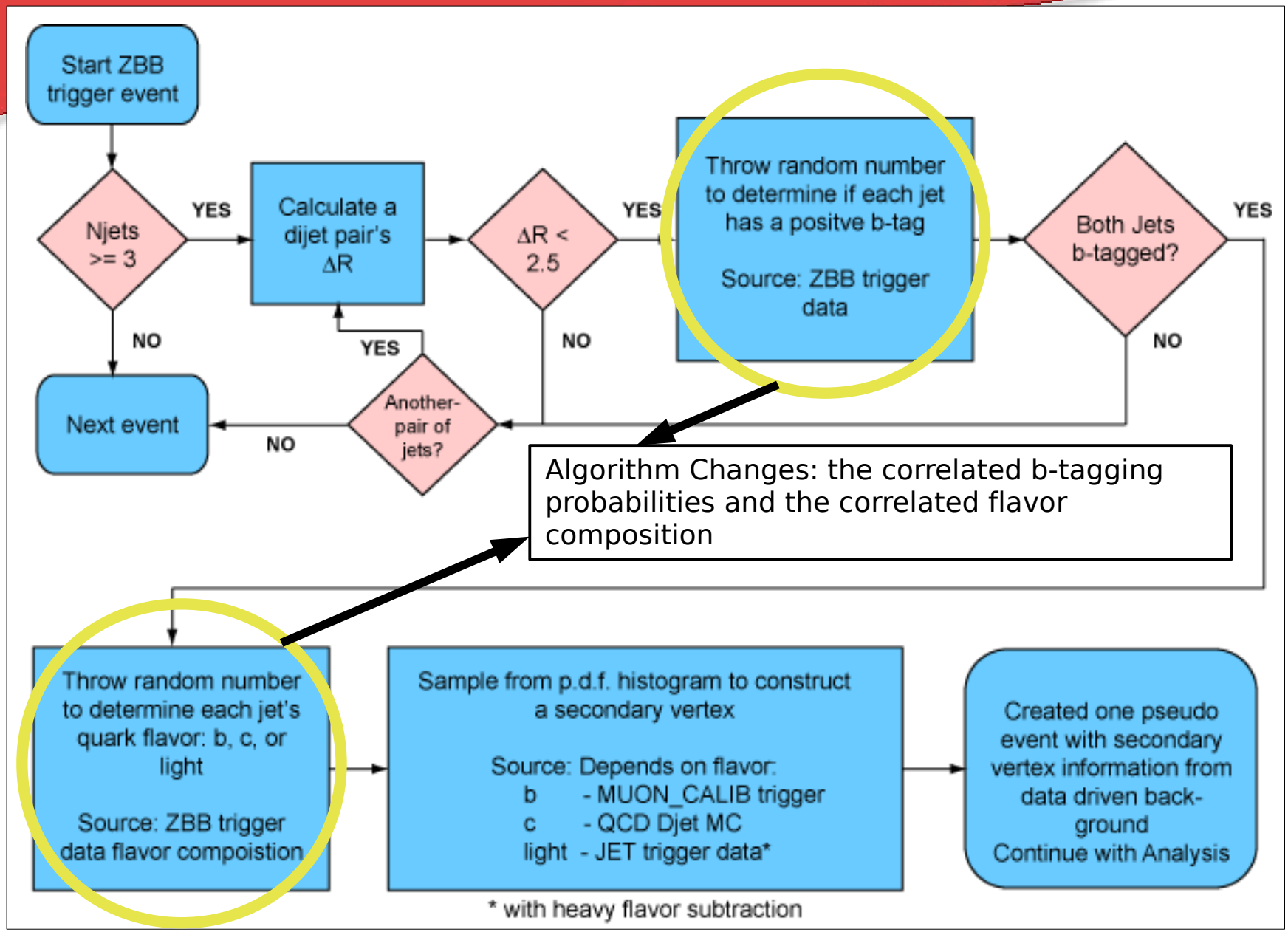
$3.5 \leq \Delta R < 4.0$

What is shown is a particular max d0 cut ($|d_0| < 0.30$ cm). This is ζ for different bins of ΔR between the two jets in the dijet pair. No difference in ζ is observed.

Flavor Correlations of Dijets

- “Does it make sense to assume the flavor of the two jets in event are not correlated ?”
  - QCD tends to produce quarks in pairs: bb, cc, etc. to preserve flavor.
- The answer is no, the flavor is correlated.
- And as a result we have to rethink how we generate the pseudo jet's flavor.
- With a dijet system, there are six possibilities for the flavors.
  - BB, BC, BL, CC, CL, and LL.
- Fit the two jets in each event simultaneously, where the vertex mass of each jet are the two variables.
  - The resulting fractions will tell what is the flavor of the dijet system, preseving the correlations between the two jet's flavors.

Flowchart



We have/will make changes to this algorithm in the following places:
1) How we generate the pseudo b-tags for the pseudo jet.
2) How we generate the flavor of each pseudo jet

Algorithm Changes

- Pseudo jet b-tag generation
  - Before: Each pseudo jet has a b-tag which is generated independently, using the probability that any jet has a b-tag (in the ZBB trigger data).
  - Future: Take into account the correlations that exists when one jet already has a b-tag.
    - The 1st jet, chosen randomly, is still assigned a b-tag in the same way
    - The other jet b-tag is generated with a different set of probabilities depending on whether the event is a control (2 jet) or signal (3 jet) event.
- Pseudo jet flavor determination
  - Before: Each pseudo jet has its flavor assigned independently using the flavor composition of the whole ZBB sample
  - Future: The pseudo dijet pair will have its flavor assigned from one of six possibilities:
    - BB, BC, BL, CC, CL, LL
    - These are determined using a simultaneous fit of two (real) jet's vertex masses.

Conclusions

- We are changing the algorithm to incorporate the feedback we received.
  - The b-tag correlations in the dijet pair will be preserved.
  - The flavor correlations in the dijet pair will be preserved.
    - Ongoing work.
- The variables in the analysis are correctly modeled in our pseudo data.
  - The shape differences that we see in the control region will be a systematic uncertainty.
- We hope to present a result soon.

Backup Slides

- ZBB Trigger Details
- Jet binning

ZBB Trigger

- Details of the trigger in the trigger table:
  - L1 :
    - one central tower with $E_T > 5$ GeV
    - two XFT tracks, $p_T^1 > 5.48$ GeV, $p_T^2 > 2.46$ GeV
  - L2 :
    - veto events w/ clusters with $E_T > 5$ GeV, $|\eta| > 1.1$
    - requires two clusters $E_T > 5$ GeV, $|\eta|<1.1$ which have $9 < \Delta\text{Wedge} < 12$
    - two SVT tracks with $p_T > 2$ GeV, $d_0 > 160$ microns, $d_0 < 1000$ microns, $\chi^2 < 12$,
      - $150 < \Delta\phi < 180$ “Opposite Side”
      - $0 < \Delta\phi < 30$ “Same Side”
      - This triggers on displaced tracks in the event.
  - L3:
    - two R=0.7 jets with $E_T > 10$ GeV, $|\eta| < 1.1$
    - two SVT tracks with $p_T > 2$ GeV, $d_0 > 160$ microns, $d_0 < 1000$ microns, $|\eta| < 1.2$
    - two tracks with $p_T > 1.5$ GeV, $d_0 > 130$ microns, $d_0 < 1000$ microns, $|\eta| < 1.2$,

      IP significance $Sd_0>3$, $\Delta z < 5$ cm
  - Dynamically Prescaled Trigger
  - This is for the latest trigger “chunk,” #17. Chunks 10-16 are nearly the same, with minor changes in the cut values, but the structure is the same.

Jet Binning

- Out jets are split into 12 different bins.
  - Four bins of E_T: [20,30) [30, 70) [70,110) [110,200)
  - Three bins of # SVT tracks: 0, 1, (>=2)
- The E_T bins are the result of the jet trigger
  - SINGLETOWER5 → [20,30)
  - JET_20 → [30,70), etc.
- The # SVT track bins are split as such because the ZBB trigger requires two SVT tracks in the event, not per jet.
  - In order to account for the differences in tag probability, flavor, and b-tag kinematics, it is necessary to split our QCD jet sample into the same SVT tracks requirements that the ZBB trigger uses.

Primary Vertex Etc

- Just to see if there is a bias due to luminosity. I split the MC into different number of primary verticies in the event.
  - Left: Signal MC and Right: QCD bb MC